SPECIFICATION DIAGNOSTICS FOR ECONOMETRIC MODELS OF DURATIONS

SUNIL SHARMA*

University of California, Los Angeles

UCLA Department of Economics
Working Paper #440
March 1987

*I would like to thank seminar participants in workshops at Brown, Cornell, UCLA and USC. I am especially grateful to Nick Kiefer, Tony Lancaster and Peter Jensen for helpful comments. The usual disclaimer applies.

ABSTRACT

This paper provides new diagnostics for evaluating the appropriateness of distributional assumptions in econometric models of duration. The exposition is in terms of the Weibull distribution, but the method is applicable more generally. The diagnostics developed are score tests of the null hypothesis of Weibull distributed durations (conditional on covariates) against alternatives which are based on expansions of various orders in a system of orthogonal functions. The statistics essentially test whether moment restrictions implied by the null specification are satisfied by the data. The paper also shows that recently developed diagnostics for uncorrected heterogeneity essentially amount to testing whether a particular moment relationship implied by the null is satisfied.

INTRODUCTION

This paper provides new diagnostics for evaluating the appropriateness of distributional assumptions in econometric models of duration. The exposition is in terms of the Weibull distribution, but the method is applicable more generally. The Weibull and exponential distributions are popular assumptions in the economic literature modelling duration of unemployment [Burdett et al (1984, 1985), Flinn and Heckman (1982), Katz (1985), Lancaster (1979), Sharma (1986)], duration of strikes and work stoppages [Horvath (1968), Newby and Winterton (1983), Kennan (1985 a,b)] and the duration of other social processes. The diagnostic statistics developed for the Weibull distribution specialize in a straightforward manner for the exponential distribution.

The diagnostics developed are score tests of the null hypothesis of Weibull distributed durations (conditional on covariates) against alternatives which are based on expansions of various orders in a system of orthogonal functions. The statistics essentially test whether moment restrictions implied by the null specification are satisfied by the data. These kinds of specification tests have been recently studied by Newey (1985) and Tauchen (1984).

Chesher (1984) developed a score test for neglected heterogeneity when the variance of the heterogeneity is small. He also showed that this score test was equivalent to the Information Matrix test proposed by White (1982). Lancaster (1985), following Cox (1983) and Chesher (1984), developed a score test for neglected heterogeneity in the context of the Weibull duration model. Sharma (1986) presented another score test for detecting neglected heterogeneity in the Weibull model. This paper shows that the approaches taken by Lancaster (1985), Kiefer (1985) and Sharma (1986) are essentially equivalent, and more importantly that the test they propose amounts to testing whether a particular moment relationship implied by the null is satisfied. In this sense the previous approaches are a special case of the method adopted here.

SECTION II

Suppose the data available to an econometrician is (t_i, X_i, δ_i), $i = 1,2,\ldots,N$ where t_i is the duration of a particular event, X_i is a (row) vector of covariates associated with the duration t_i, and δ_i is an indicator variable depending on whether the duration is censored or not. An econometric duration model attempts to find the effect of covariates on the duration of an event, and also whether there is any duration dependence. Further, the econometrician would like to know whether conditional on covariates the distribution of t_i can be reasonably approximated by some simple distribution.

Most duration models are specified in terms of the hazard function, which is merely a convenient transformation of the density function.[1] The hazard function of a distribution with density function f(t) and distribution function F(t) is

$$\psi(t) = \lim_{\Delta t \to 0} \frac{\text{prob}\,(t < T < t + \Delta t)}{\Delta t} \cdot \frac{1}{\text{prob}\,(T > t)}$$

$$= \frac{f(t)}{1 - F(t)} = \frac{f(t)}{\bar{F}(t)} \qquad (1)$$

where $\bar{F}(t) = 1 - F(t)$ is called the survivor function. The density and survivor functions can be written in terms of the hazard function as

$$f(t) = \psi(t) \exp\left\{ -\int_0^t \psi(u)\, du \right\}$$

$$\bar{F}(t) = \exp\left\{ -\int_0^t \psi(u)\, du \right\} \qquad (2)$$

---

[1] see Cox (1962).

When we have covariates the conditional hazard function can be written as

$$\pi(t,x) = \psi(t) \cdot \phi(x) \qquad \text{—— (3)}$$

The choice of $\phi(t)$ depends on the data. A natural assumption for $\phi(x)$ is $\exp\{x\beta\}$, where β is a column vector of coefficients, because it guarantees the non-negativity of $\pi(t,x)$. Other specific forms for $\phi(X)$ can be used,[2] but most suffer from the disadvantage that the set of β values have to be restricted to ensure non-negativity of the hazard and, also because such restrictions add to the complexity of computation.

Consider the hazard model

$$\pi(t,X) = \psi(t) \ \exp(x\beta) \qquad \text{—— (4)}$$

This specialization of (3) is called the proportional hazard model and is a widely used framework for specifying duration models.[3] The exposition in this paper uses the Weibull distribution as the null case, implying

$$\psi(t) = \alpha t^{\alpha-1} \qquad \text{—— (5)}$$

The Weibull distribution is a reasonably flexible assumption allowing for both positive and negative duration dependence. The Weibull distribution and its specialization the exponential distribution are popular assumptions in econometric duration models. [See for example Burdett et al (1980, 1984, 1985), Flinn and Heckman (1982), Newby and Winterton (1983), Katz 1985), Kennan (1985a,b), Lancaster (1979), Sharma (1986)] However, it should be

---

[2]see Feigl and Zelen (1965), Zippan and Armitage (1966), Greenberg et al (1974).

[3]The proportional hazard framework was proposed by Cox (1972). A review of the framework is provided by Breslow (1975), Holford (1976) and Kay (1977).

emphasized that the method of developing diagnostics presented here can be used for arbitrary null distributions.

The hazard, density and survivor function conditional on covariates are written as

$$\begin{aligned} \pi(t,X) &= \alpha t^{\alpha-1} \exp\{X\beta\} \\ f(t,X) &= \alpha t^{\alpha-1} \exp\{X\beta\} \exp\{-t^{\alpha} \exp(X\beta)\} \\ \bar{F}(t,X) &= \exp\{X\beta - t^{\alpha} \exp(X\beta)\} \end{aligned} \qquad (6)$$

Suppose the data comprise of N_1 complete spells and N_2 right censored durations. Under the assumption that the observations are independent the loglikelihood function is

$$l = \sum_{i=1}^{N_1} \{\ln\alpha + (\alpha-1)\ln t_i + X_i\beta - t_i^{\alpha} \exp(X_i\beta)\} + \sum_{j=1}^{N_2} \{-t_j^{\alpha} \exp(X_j\beta)\}$$

$$l = \sum_{i=1}^{N_1} \{\ln\alpha + (\alpha-1)\ln t_i + X_i\beta\} + \sum_{j=1}^{N} \{-t_j^{\alpha} \exp(X_j\beta)\}$$

$$l = \sum_{i=1}^{N} \left[\delta_i \{\ln\alpha + (\alpha-1)\ln t_i + X_i\beta\} - \{t_i^{\alpha} \exp(X_i\beta)\}\right] \qquad (7)$$

$$\text{where } \delta_i = \begin{cases} 1 & \text{if spell is complete} \\ 0 & \text{if spell is censored} \end{cases}$$

Estimates for the parameters $[\alpha\ \beta']'$ are obtained by maximizing (7). Let $X = [X_1,\ldots,X_m]$ be a 1xm vector with the first term being one and the rest representing (m-1) covariates. Let $\gamma = [\alpha\ \beta']'$ be a (m+1)x1 vector of parameters, where β is a mx1 vector X. The first and second derivatives of l with respect to parameter vector γ are

$$\frac{\delta l}{\delta \gamma} = \begin{bmatrix} \frac{\delta l}{\delta \alpha} \\ \frac{\delta l}{\delta \beta_1} \\ | \\ | \\ \frac{\delta l}{\delta \beta_m} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{N} \left[\delta_i \left(\frac{1}{\alpha} + \ln t_i \right) - \left(t_i \left(\ln t_i \right) \exp \left(X_i \beta \right) \right) \right] \\ \sum_{i=1}^{N} \left[\delta_i X_{i1} - t_i^{\alpha} \exp \left(X_i \beta \right) X_{i1} \right] \\ | \\ | \\ \sum_{i=1}^{N} \left[\delta_i X_{im} - t_i^{\alpha} \exp \left(X_i \beta \right) X_{im} \right] \end{bmatrix} \quad \text{—— (8)}$$

$$\frac{\delta^2 l}{\delta \gamma' \delta \gamma'} = \begin{bmatrix} \frac{\delta^2 l}{\delta \alpha^2} & \frac{\delta^2 l}{\delta \alpha \delta \beta'} \\ \frac{\delta^2 l}{\delta \alpha \delta \beta} & \frac{\delta^2 l}{\delta \beta \delta \beta'} \end{bmatrix} \quad \text{—— (9)}$$

where

$$\frac{\delta^2 l}{\delta \alpha^2} = \sum_{i=1}^{N} \left[\delta_i \left(- \frac{1}{\alpha^2} \right) - \left(t_i^{\alpha} \left(\ln t_i \right)^2 \exp \left(X_i \beta \right) \right) \right]$$

$$\frac{\delta^2 l}{\delta \alpha \delta \beta'} = \sum_{i=1}^{N} - t_i^{\alpha} \left(\ln t_i \right) \exp \left(X_i \beta \right) . X_i$$

$$\frac{\delta^2 l}{\delta \beta \delta \beta'} = \sum_{i=1}^{N} - t_i^{\alpha} \exp \left(X_i \beta \right) . X_i' X_i$$

The maximum likelihood equations $\frac{\delta l}{\delta \gamma} = 0$ can be solved for $\hat{\gamma}$, the maximum likelihood parameter estimates, using the Newton-Raphson method.[4]

[4]The loglikelihood function is concave in $\gamma = [\alpha \ \beta']'$ over its entire domain. Hence, the Newton-Raphson method leads to a global maximum, see Sharma (1986).

Let $I(\gamma) = E\left[-\frac{\delta^2 l}{\delta\gamma\delta\gamma'}\right]$ be the information matrix. For the central limit theorem to apply to the score $\frac{\delta l}{\delta\gamma}$, we require certain restrictions on covariate vectors X_i and the censoring mechanism. The relative information from any observation i, should approach the zero matrix as $N \to \infty$. Also, the censoring times must not converge too rapidly to zero, as $N \to \infty$. Given these conditions, the asymptotic distribution of the score $\frac{\delta l}{\delta\gamma}$ is multivariate (m+1) normal with mean zero and variance $I(\gamma)$, whereas that of $\hat{\gamma}$ is multivariate (m+1) normal with mean vector γ and variance $I^{-1}(\gamma)$. In both asymptotic results, $I(\gamma)$ can be replaced by the observed information matrix $I(\gamma) = \left[-\frac{\delta^2 l}{\delta\gamma\delta\gamma'}\right]_{\gamma = \hat{\gamma}}$.

SECTION III

In this section, diagnostics for assessing the appropriateness of the Weibull specification are developed. The statistics are Lagrange multiplier or score tests[5] of the null hypothesis of Weibull distributed durations (conditional on regressors) against alternatives in a family of approximations to arbitrary distributions of non-negative random variables. The alternatives considered are defined using expansions of various orders in a system of orthogonal functions (which are generalizations of Laguerre polynomials).[6] First, the system of orthogonal functions used is defined, then the class of alternative hypotheses is developed and lastly the diagnostic statistics are derived.

We define a system (sequence) of functions $\{f_n(n),\ n = 0,1,2,\ldots\}$ to be orthogonal on the interval $a \le t \le b$, with respect to the weight function $w(t)$, if

$$\int_a^b f_n(t)\ f_m(t)\ w(t)\ dt = 0 \qquad \text{—— (10)}$$

$$\text{for } n \neq m\ ;\quad n,m = 0,1,2,\ldots\ldots$$

The system of functions we consider are defined by

$$L_n(t) = \sum_{r=0}^{N} \binom{n}{n-r} \frac{(-t^{\alpha})^r}{r!} \qquad \text{—— (11)}$$

$$n = 0,1,2,\ldots\ldots$$

---

[5]see Breusch and Pagan (1980), and Engle (1982).

[6]see Szego (1975), Chihara (1978).

where $\alpha > 0$ is a parameter.[7] The weighting function is

$$w(t) = \alpha\ t^{\alpha-1}\ \exp\ (-t^{\alpha}) \qquad (12)$$

The following theorem is basic to the development and interpretation of the diagnostic statistics proposed.

<u>THEOREM</u>: The family of functions $\{L_n(t)\}$ defined in (11) is orthogonal on the interval $0 \leq t \leq \infty$, with respect to the Weibull weighting function defined in (12), that is

$$\int_0^{\infty} L_n(t)\ L_m(t)\ \alpha\ t^{\alpha-1}\ \exp\ (-t^{\alpha})\ dt = \delta_{mn}$$

$$\text{where } \delta_{mn} = \begin{cases} 1 & \text{if } m = n \\ 0 & \text{if } m \neq n \end{cases} \qquad (13)$$

<u>Proof</u>: see Appendix A.

The first five functions in the family $\{L_n(t)\}$ are

$$L_0\ (t) = 1$$

$$L_1\ (t) = 1 - t^{\alpha} \qquad (14)$$

$$L_2\ (t) = \frac{1}{2}\left[\ t^{2\alpha} - 4t^{\alpha} + 2\ \right]$$

$$L_3\ (t) = \frac{1}{6}\left[\ -\ t^{3\alpha} + 9t^{2\alpha} - 18t^{\alpha} + 6\ \right]$$

$$L_4\ (t) = \frac{1}{24}\left[\ t^{4\alpha} - 16t^{3\alpha} + 72t^{2\alpha} - 96t^{\alpha} + 24\ \right]$$

The family of alternatives to the Weibull defined below is one in which the Weibull distribution is nested and one which allows a number of interesting departures from the Weibull. Let $p(t)$ be the density of a non-negative random variable (eg. duration of a particular event). Since the

---

[7] Note that for $\alpha = 1$, the system $\{L_n\ (t)\}$ reduces to the Laguerre polynomials.

density is unknown, we assume $p(t)$ belongs to (or can be approximated) by the family

$$p(t) = \alpha t^{\alpha-1} \exp(-t^{\alpha}) \left[1 + \sum_{j=1}^{\infty} \theta_j L_j(t) \right] \qquad (15)$$

where $L_j(t)$ is the j^{th} member in the orthogonal family of functions $\{L_n(t)\}$. The Weibull distribution is nested in the family since $p(t) = \alpha t^{\alpha-1} \exp(-t^{\alpha})$ when the θ_j are zero for all j. Hence, testing the null hypothesis that the distribution is Weibull reduces to testing whether $\theta_j = 0$ for all j.

The testing of $\theta_j = 0$ has a simple interpretation. Multiplying both sides of (15) by $L_k(t)$ and integrating gives

$$E[L_k(t)] = \int_0^{\infty} L_k(t) \; p(t) \; dt$$

$$= \int_0^{\infty} L_k(t) \; L_0(t) \; \alpha t^{\alpha-1} \exp(-t^{\alpha}) \; dt$$

$$+ \sum_{j=1}^{\infty} \theta_j \int_0^{\infty} L_k(t) \; L_j(t) \; \alpha t^{\alpha-1} \exp(-t^{\alpha}) \, dt$$

$$= \theta_k \qquad (16)$$

For example, testing $\theta_3 = 0$ essentially amounts to asking whether the data satisfies the following moment condition

$$E[-t^{3\alpha} + 9t^{2\alpha} - 18t^{\alpha} + 6] = 0 \qquad (17)$$

Hence, the score tests analyzed here are essentially specification tests based on moment conditions like (17). This class of specification tests, called M-tests, have been studied recently by Newey (1985) and Tauchen (1984).

In practice $p(t)$ can be approximated by

$$p(t) = \alpha t^{\alpha-1} \exp(-t^{\alpha}) \; [1 + \theta_1 L_1(t) + \ldots + \theta_n L_n(t)] \qquad \text{—— (18)}$$

which represents a truncation of (15). It is possible that for the unknown density that $\theta_j = 0,\ j = 2, \ldots, n$ there are θ_k for $k > n$ which are significantly different from zero. Clearly, this is not a practical problem because if a distribution agrees with the Weibull in the first four or five moments it can be reasonably assumed to be Weibull without any great loss.[8]

We now derive the diagnostic statistics. It should be emphasized here that the hypothesis being tested is examined conditional on estimated values of the parameters $\gamma = [\alpha\ \beta']'$. The question addressed is whether conditional on the best estimates (by the maximum likelihood criterion), there is any evidence of misspecification.[9] In what follows it should be understood that α and β are the maximum likelihood estimates of the parameters estimated under the null. The null hypothesis is that the duration distribution (conditional on regressors) is specified by

$$f(t;X) = \alpha t^{\alpha-1} \exp(X\beta) \exp(-t^{\alpha} e^{X\beta}) = \alpha t^{\alpha-1} \eta \exp(-\eta t^{\alpha}) \qquad \text{—— (19)}$$

where $\eta = \exp(X\beta)$.

Using the transformation $z^{\alpha} = \eta t^{\alpha}$ we see that z has a standard Weibull density

$$p(z) = \alpha z^{\alpha-1} \exp(-z^{\alpha}) \qquad \text{—— (20)}$$

The class of alternatives is defined by

$$p^*(z) = \alpha z^{\alpha-1} \exp(-z^{\alpha}) \; [1 + \theta_1 L_1(z) + \theta_2 L_2(z) + \ldots + \theta_n L_n(z)] \qquad \text{—— (21)}$$

---

[8] For larger sample size the researcher may want to consider a larger class of alternatives by increasing n in (15).

[9] see Kiefer (1985).

The above density cannot be directly used since z contains unknown parameters. Changing variables, (using the transformation $z = \eta^{1/\alpha} t$) defines the class of alternatives as

$$p^*(t,X) = \alpha t^{\alpha-1} \eta \exp(-\eta t^{\alpha}) [1 + \theta_1 L_1 (\eta^{1/\alpha} t) + \theta_2 L_2 (\eta^{1/\alpha} t) + \ldots + \theta_n L_n (\eta^{1/\alpha} t) \qquad (22)$$

Consider the case when all the observed durations are complete. [The case when there are complete and censored observations is considered in Appendix B]. Using (22), N^{-1} times the loglikelihood function under independent sampling is

$$l = \frac{1}{N} \sum_{i=1}^{N} \ln p_i^* (t,X_i) \qquad (23)$$

$$= \ln \alpha + N^{-1}(\alpha - 1) \sum_{i=1}^{N} \ln (t_i) - N^{-1} \sum_{i=1}^{N} X_i\beta$$

$$- N^{-1} \sum_{i=1}^{N} t_i^{\alpha} e^{x_i\beta} + N^{-1} \sum_{i=1}^{N} \ln C_i$$

$$\text{where } C_i = [1 + \theta_1 L_1 (\eta^{\frac{1}{\alpha}} t) + \ldots + \theta_n L_n (\eta^{\frac{1}{\alpha}} t)]$$

The null hypothesis is that conditional on X, the duration distribution is Weibull. This is equivalent to asking whether $\theta_j = 0$, $j = 1,2,\ldots,n$. The score vector under the null is given by

$$\frac{\delta l}{\delta \theta} = \begin{bmatrix} \frac{\delta l}{\delta \theta_1} \\ \frac{\delta l}{\delta \theta_2} \\ | \\ | \\ \frac{\delta l}{\delta \theta_n} \end{bmatrix} = \begin{bmatrix} \frac{1}{N} \sum_{i=1}^{N} L_1 \; (\eta_i^{\frac{1}{\alpha}} t_i) \\ \frac{1}{N} \sum_{i=1}^{N} L_2 \; (\eta_i^{\frac{1}{\alpha}} t_i) \\ | \\ | \\ \frac{1}{N} \sum_{i=1}^{N} L_n \; (\eta_i^{\frac{1}{\alpha}} t_i) \end{bmatrix} \qquad \text{—— (24)}$$

Notice that $\left.\frac{\delta l}{\delta \theta_j}\right|_{\theta = 0}$ is just the sample mean of the j^{th} function in the orthogonal system of functions $\{L_n (t)\}$. Also, note that the test of $\theta_1 = 0$ is not meaningful because

$$\left.\frac{\delta l}{\delta \theta_1}\right| = N^{-1} \sum_{i=1}^{N} (1 - t_i^{\alpha} \; e^{X_i \beta}) \qquad \text{—— (25)}$$

is identically zero when evaluated at the maximum likelihood estimates of α and β.

Consider first the testing of the simple hypothesis $\theta_2 = 0$. The Lagrange multiplier statistic is given by

$$Z_{(2)} = \frac{\left\{ \frac{\delta l}{\delta \theta_2} \right\}^2}{V_{(2)}} \qquad \text{—— (26)}$$

In (26), $V_{(2)}$, an estimate of the variance of $\frac{\delta l}{\delta \theta_2}$, is given by

$$V_{(2)} = \frac{1}{N^2} \sum_{i=1}^{N} [L_{2i} - \bar{L}_{2i}]^2 \qquad \text{—— (27)}$$

where $L_{2i} = L_2 (\eta_i^{\frac{1}{\alpha}} t)$ and the bar denotes the sample mean. The proposed diagnostic is a score test where the variance of the score is not directly based on the information matrix. From standard score theory, $Z_{(2)}$ has an asymptotic distribution which is χ^2 (1), and the calculated value of $Z_{(2)}$ is compared with the appropriate significance points of the distribution.

In order to test $\theta_2 = \theta_3 = \ldots = \theta_n = 0$ we use the score vector defined in (24) and an estimate of its variance, the (n-1) x (n-1) matrix $V_{(n)}$, whose (j,k) term is given by

$$V_{(n)} (j,k) = \frac{1}{N^2} \sum_{i=1}^{N} (L_{ji} - \bar{L}_{ji}) (L_{ki} - \bar{L}_{ki}) \qquad \text{—— (28)}$$

The statistic used is $Z_{(n)}$, defined by

$$Z_{(n)} = \left[\frac{\delta l}{\delta \theta} \right] \left[V_{(n)} \right]^{-1} \left[\frac{\delta l}{\delta \theta} \right] \qquad \text{—— (29)}$$

which has an asymptotic χ^2 (n-1) distribution.

SECTION IV

In this section we consider score tests for heterogeneity. Heterogeneity is defined as unmeasured and measured exogenous variables that vary across individuals in a particular sample. In duration models not correcting for heterogeneity across observed units in a sample can lead to biased estimates of duration dependence and coefficients of the included covariates, and hence have serious consequences for the interpretation of econometric results.

Suppose that the individual characteristics in the 1xm covariate vector X do not sufficiently control for heterogeneity or that the presence of "unobservables" does not justify the assumption of homogeneity for the sample under consideration. In the Weibull model the actual hazard function is

$$\pi = \alpha t^{\alpha-1} \exp (X\beta + \nu) \qquad \text{—— (30)}$$

where ν is an unobserved random variable whose distribution is not known. There are three possible interpretations of ν.[10] We assume that vector X contains a constant term and therefore $X\beta$ can be written as $\beta_1 + X^{(1)}\beta^{(1)}$, when $X^{(1)} = (X_2, X_3, \ldots, X_m)$ and $\beta^{(1)} = (\beta_2, \beta_2, \ldots, \beta_m)$. We can interpret ν as heterogeneity with respect to the parameter β_1. The other two interpretations arise out of the effects of measurement errors in the vector $X^{(1)}$ or duration t.

Let the distribution of ν across the observed units be given by the density $q(\nu)$. The duration density of t, given X <u>and</u> ν is

$$f(t,X,\nu) = \alpha t^{\alpha-1} \exp (X\beta + \nu) \exp \{-t^{\alpha} \exp (X\beta + \nu)\} \qquad \text{—— (31)}$$

---

[10] see Lancaster (1985).

while the duration density conditional on X alone, and which is the density relevant for data analysis is

$$f(t;\ X) = \int_{-\infty}^{\infty} f\ (t;\ X,\ \nu)\ \ q\ (\nu)\ \ d\nu. \qquad \text{—— (32)}$$

Note if ν was a degenerate random variable (i.e., there was no uncorrected heterogeneity), (32) would be identical to (19).

In the simple linear regression model, omitted regressors create a problem only if the left-out variables are correlated with those included. However, in the duration models we are considering, this no longer holds. To see this formally, assume that ν is distributed independently of X and t. The survivor function conditional on X and ν is

$$\bar{F}(t;X,\nu) = \exp\ \{-t^{\alpha}\ \exp\ (X\beta + \nu)\} \qquad \text{—— (33)}$$

and that conditional on X alone is

$$\bar{F}\ (t;X) = \int_{-\infty}^{\infty} \exp\ \{-t^{\alpha}\ \exp\ (X\beta + \nu)\}\ \ q(\nu)\ d\nu \qquad \text{—— (34)}$$

The hazard function of the survivor function in (34) can be derived as

$$\pi\ (t;X) = -\frac{d}{dt}\ \ln\ \bar{F}\ \ (t,X) \qquad \text{—— (35)}$$

$$= \alpha t^{\alpha-1}\ e^{X\beta} \int_{-\infty}^{\infty} e^{\nu}\ .\ \ \phi\ (\nu;x,t)\ \ d\nu$$

$$\text{where}\quad \phi\ (\nu;x,t) = \frac{\exp\ \ \{-t^{\alpha}\ \ \exp\ (X\beta + \nu)\}\ \ q(\nu)}{\int_{-\infty}^{\infty} \exp\ \{-t^{\alpha}\ \ \exp\ (X\beta + \nu)\}\ \ q(\nu)d\nu}$$

is the density of ν, conditional on duration greater than t.

If there are no omitted regressors and ν is known, we get from (30)

$$\ln \pi(t;X,\nu) = \ln\alpha + (\alpha - 1) \ln t + X\beta + \nu \qquad (36)$$

$$\frac{\delta \ln \pi(t;X,\nu)}{\delta x_j} = \beta_j \qquad (37)$$

If x_j is the natural logarithm of some economic variable, β_j can be interpreted as an elasticity with respect to that variable. With omitted covariate we have, from (35) that

$$\ln \pi(t,X) = \ln\alpha + (\alpha - 1) \ln t + X\beta + \ln E\,(e^{\nu} \mid \tilde{t} > t) \qquad (38)$$

$$\frac{\delta \ln \pi(t,X)}{\delta X_j} = \beta_j\,[1 - t^{\alpha}\ \exp(X\beta)\ \text{Var}\,(e^{\nu} \mid \tilde{t} > t)] \qquad (39)$$

This implies that the impact of included economic variables on the hazard rate, besides being diminished is also dependent on the duration t. Hence, the estimates of the coefficients of the included covariates can be biased even though the excluded variables are uncorrelated with them. Note if there is no heterogeneity with respect to ν, (39) reduces to (37).

The standard practice in recent econometric work has been to assume that $f(t;X,\nu)$ and $q(\nu)$ belong to some simple parametric families of distributions. The choice of the mixing distribution $q(\nu)$ is generally arbitrary, being justified in terms of analytical or computational convenience. Since the estimates of the parameters of duration models maybe sensitive to choice of the mixing distribution it is important to have diagnostics which point out the extent of heterogeneity.[11] Further, since $q(\nu)$ is generally unknown, we need a statistic which does not require any particular specification of $q(\nu)$.

Given a constant term in vector X, without loss of generality, we can assume $E(\nu) = 0$. Also, let the variance of ν be σ^2. The null hypothesis

[11] see Heckman and Singer (1985).

is that there is no heterogeneity which is equivalent to testing whether $\sigma^2 = 0$. Using the integrated hazard

$$\epsilon\ (t;X) = \int_{-\infty}^{\infty} \pi(s,X,\nu = 0)\ ds \qquad (40)$$

we write the density function as

$$f(t;X,\nu) = \pi(t,X,0)\ e^{\nu}\ \exp\,[-\epsilon(t;X)\ e^{\nu}] \qquad (41)$$

Assuming that the heterogeneity in the data is small, Kiefer (1984) and Sharma (1986) use a second order Taylor series around $\nu = 0$ to approximate $f(t;X,\nu)$.

$$f(t;X,\nu) = f(t,X,0) + \nu \left[\frac{\delta f(t;X,0)}{\delta\nu} \right]_{\nu=0} + \frac{\nu^2}{2} \left[\frac{\delta^2 f\ (t;X,\nu)}{\delta\nu^2} \right]_{\nu=0} \qquad (42)$$

Integrating above equation with respect to ν and using $E(\nu) = 0$ gives the approximate density

$$f(t;X) \approx f(t;X,0) + \frac{\sigma^2}{2} \left[\frac{\delta^2 f(t;X,\nu)}{\delta\nu^2} \right]_{\nu=0} \qquad (43)$$

$$= f^*\ (t;X)$$

It is easy to show that

$$\left[\frac{\delta^2 f(t;X,\nu)}{\delta\nu^2} \right]_{\nu=0} = f(t;X,0)\ \ [1 - 3\ \epsilon(t;X)e^{\nu} + \epsilon^2\ (t;X)\ e^{2\nu}] \qquad (44)$$

Substituting in (43) we get

$$f^*(t;X) = f(t;X,0)\ \left[1 + \frac{\sigma^2}{2}\ (1 - 3\ \epsilon\ e^{\nu} + \epsilon^2\ e^{2\nu}) \right] \qquad (45)$$

The proposed diagnostic for heterogeneity is a score statistic for the hypothesis that $\sigma^2 = 0$, based on the approximate density $f*(t,X)$ given in (43). Consider the case when all the durations are complete. (The censored case is considered in Appendix C). Defining $l*$ to be the loglikelihood based on $f*(t;X)$, the score for σ^2 is

$$\frac{\delta l*}{\delta \sigma^2} = \frac{1}{2} \sum_{i=1}^{N} \left[\frac{1 - 3\,\epsilon_i + \epsilon_i^{\,2}}{1 + \frac{\sigma^2}{2}\ (1 - 3\,\epsilon_i + \epsilon_i^{\,2}\ ))} \right]$$

where $\epsilon_i = \epsilon(t_i,\ X_i)$. —— (46)

The mean score under the null is

$$S = \frac{1}{2N} \sum_{i=1}^{N} (1 - 3\,\epsilon_i + \epsilon_i^2\ ) \qquad \text{—— (47)}$$

The diagnostic is a standard test of the hypothesis that $E(S) = 0$ against $E(S) > 0$. The test statistic is

$$Z = \frac{S}{\sqrt{Var(S)}} \qquad \text{—— (48)}$$

which under the null has an asymptotic distribution which is standard normal.

Lancaster (1985), following Cox (1983) and Chesher (1984) developed another test for neglected heterogeneity. We show that the approach taken by Lancaster and that by Kiefer (1984) and Sharma (1986) lead to the same statistic under the null. Lancaster writes the hazard function conditional on heterogeneity term u as

$$\bar{F}\ (t;X,u) = \exp\ (-\epsilon.u) \qquad \text{—— (49)}$$

where $E(u) = 1$ and variance of u is σ^2. The unconditional hazard function is

$$\bar{F}*(t;X) = E_u\ [\exp\ (-\epsilon.u)] = \int \exp\ (-\epsilon\ u)\ .\ q(u)\ du.$$

Approximating $\exp(-\epsilon.u)$ by a second-order Taylor expansion around $u = 1$ we obtain

$$\bar{F}^*(t;X) = E_u \left[e^{-u} - (u - 1)\, e^{-\epsilon}\, \epsilon + (u - 1)^2\, e^{-\epsilon}\, \frac{\epsilon^2}{2} \right]$$

$$= \exp(-\epsilon) \left[1 + \sigma^2 \frac{\epsilon^2}{2} \right]$$

$$= \bar{F}\,(t;X,u=1) \left[1 + \sigma^2 \frac{\epsilon^2}{2} \right] \qquad \text{—— (51)}$$

The density based on the "approximate" distribution in (51) is

$$f^*(t;X) = f(t,X,u=1) \left[1 + \frac{\sigma^2}{2} (\epsilon^2 - 2\epsilon) \right] \qquad \text{—— (52)}$$

Lancaster considers testing for neglected heterogeneity by a score test of $\sigma^2 = 0$ in the family defined in (52).

The score for σ^2, based on N completed durations[12] is

$$\frac{\delta l^*}{\delta\sigma^2} = \frac{1}{2} \sum_{i=1}^{N} \frac{\left[\epsilon_i^2 - 2\epsilon_i \right]}{\left[1 + \frac{\sigma^2}{2} (\epsilon_i^2 - 2\epsilon_i) \right]} \qquad \text{—— (53)}$$

where l^* is the loglikelihood based on the density in (52). The mean score under the null is

$$\frac{\delta l^*}{\delta\sigma^2} = \frac{1}{2N} \sum_{i=1}^{N} \left[\epsilon_i^2 - 2\epsilon_i \right] \qquad \text{—— (54)}$$

Since the integrated hazard ϵ has a unit exponential distribution, the scores derived in (47) and (54) are essentially the same.

We now show that in the Weibull model the above tests of neglected heterogeneity are equivalent to testing $\theta_2 = 0$ in the family defined by

---

[12] The censored case is discussed in Appendix C.

(22). The score element for testing $\theta_2 = 0$ is

$$\frac{\delta l}{\delta \theta_2} = \frac{1}{N} \sum_{i=1}^{N} L_2 \; (\eta_i^{\frac{1}{\alpha}} t_i)$$

$$= \frac{1}{N} \sum_{i=1}^{N} \frac{1}{2} \left[t_i^{2\alpha} \eta_i^2 - 4 \; t_i^{\alpha} \; \eta_i + 2 \right]$$

$$= \frac{1}{2N} \sum_{i=1}^{N} \left[\epsilon_i^2 - 4 \epsilon_i + 2 \right] \qquad \text{——— (55)}$$

where $\epsilon_i = t_i^{\alpha} \eta_i = t_i^{\alpha} \exp (X_i \beta)$ is the integrated hazard in the Weibull model. Again since ϵ has a unit exponential distribution it is easy to see that (47), (54) and (55) are essentially testing the same moment restriction.

SECTION V

In this section we give another interpretation to the statistics developed in sections III and IV. Let $l = \ln f(t;X)$ be the loglikelihood function of a duration distribution (conditional on covariates X) depending on scalar parameter β and let $l_j = \frac{\delta^j l}{\delta \beta^j}$. Then, by repeatedly differentiating both sides of

$$\int \exp\{l\}\, dt = 1 \qquad \text{—— (56)}$$

we get the sequence of identities

$$E[\, l_2 + l_1^2 \,] = 0 \qquad \text{—— (57)}$$

$$E[\, l_3 + 3 l_1 l_2 + l_1^3 \,] = 0 \qquad \text{—— (58)}$$

and so on.

If $f(t,X)$ is the Weibull duration distribution conditional on covariates defined by

$$f(t;X) = \alpha t^{\alpha-1} e^{X\beta} \exp\{-t^{\alpha} e^{X\beta}\} \qquad \text{—— (59)}$$

(where the first element of X is one with corresponding coefficient β_0.)

and $l_j = \frac{\delta^j l}{\delta \beta_0^j}$, then the moment restriction implied by (57) is

$$E[\, 1-3\epsilon + \epsilon^2 \,] = 0 \qquad \text{—— (60)}$$

and that implied by (58) is

$$E[\, 1-7\epsilon + 6\epsilon^2 - \epsilon^3 \,] = 0 \qquad \text{—— (61)}$$

The moment restriction (60) is essentially the same as that for testing $\theta_2 = 0$. Also, since $E(\epsilon) = 1$ and $E(\epsilon^2) = 2$, it is easy to see that asking whether the data satisfies the restriction implied by (61) is the same as testing $\theta_3 = 0$ in the family defined by (22).

The above analysis makes clear that the tests developed in section III have an interpretation as M-tests in the sense defined by Newey (1985).

SECTION VI

The data used to illustrate the use of the statistics developed is from the Denver Income Maintenance Experiment (DIME). The purpose of this experiment was to measure the effect of a negative income tax on labor supply. The families enrolled in DIME were followed for forty-eight months and met the following criteria

(i) race: head of family had to be white, black or hispanic

(ii) family type: a two-head family or a single head, with at least one dependent.

(iii) family income: pre-experiment earnings in (1970-71 dollars) were under $9,000 for a family of four with one head, or under $11,000 for a two-head family.

(iv) head of family: age 18 to 58 years and capable of gainful employment.

The data were used to construct histories of labor market status (number, timing and sequence of all change in labor market status) using Current Population Survey definitions. The unit of observation was the length of a spell of employment or unemployment. Complete and censored spells were differentiated and for each spell a number of individual characteristics and labor market variables relating to the person experiencing the spell were recorded.[13]

We consider spells of employment and unemployment (non-employment) which were initiated during the period of participation in the sample. For each individual, one spell the first is included. If the spell concludes by

---

[13]Lundberg (1981) has discussed the construction of the data on labor market spells. Also see Burdett et al (1984) and Sharma (1986).

a transition to the other state it is called complete, else it is called (right) censored. Table 1 gives summary statistics of the data used. The covariates used in the analysis to correct for heterogeneity are race, education and age. The hazard function specification and estimation was detailed in Section II.

Maximum likelihood estimates of the employment hazard function under the Weibull assumption are given in Table 2. The Weibull specification for the full sample cannot be rejected at the 5% significance level using test statistics $Z_{(2)}$ and $Z_{(3)}$. However, $Z_{(4)}$ does reject the Weibull specification at the same level of significance. The hazard estimates by race category are reported in the last three columns of Table 2. The Weibull specification is clearly acceptable for the white sample and does a reasonable job for the Black sample. However, for the Hispanic sample although $Z_{(2)}$ and $Z_{(3)}$ do not reject the null, $Z_{(4)}$ decisively does so. The breakdown of the Weibull specification for the Hispanic sample occurs in the fourth moment. This brings out the importance (especially in large samples) of examining a wider class of alternatives by considering a larger number of terms of the system of orthogonal functions when defining the family in which the null is embedded. It should be mentioned here that density approximations based on orthogonal functions have the appeal that under the null higher order terms are uncorrelated with lower order terms, and hence neglecting higher order terms required for an exact specification of the class of alternatives may not lead to serious problems in estimating lower order terms.

Table 3 and 4 gives estimates for the unemployment hazard under the Weibull and Exponential assumptions respectively. The Exponential specification is rejected for the full sample and the White sample. It is

TABLE 1

SUMMARY STATISTICS

EMPLOYMENT SPELLS INITIATED DURING PERIOD OF PARTICIPATION IN SAMPLE. ONE SPELL PER INDIVIDUAL, THE FIRST, IS INCLUDED.

	Full Sample (N Observations = 519)		Complete Spells (N Observations = 317)	
Variable	Mean	Standard Deviation	Mean	Standard Deviation
Duration (weeks)	52.40	60.26	37.47	40.95
Age (Years)	29.37	10.02	27.68	9.32
Education (grade completed	11.09	2.13	11.16	2.13
Black (proportion)	0.337	....	0.334	....
Hispanic	0.324	....	0.353	....

UNEMPLOYMENT SPELLS INITIATED DURING PERIOD OF PARTICIPATION IN SAMPLE. ONE SPELL PER INDIVIDUAL, THE FIRST, IS INCLUDED.

	Full Sample (N Observations = 348)		Complete Spells (N Observations = 256)	
Variable	Mean	Standard Deviation	Mean	Standard Deviation
Duration (weeks)	14.2	17.1	11.4	12.5
Age (Year in 1970)	29.3	9.96	28.8	9.31
Education (grade completed	11.0	2.16	11.0	2.21
Black (proportion)	0.327	....	0.313	....
Hispanic	0.348	....	0.328	....

TABLE 2

HAZARD FUNCTION ESTIMATES AND TEST STATISTICS.
EMPLOYMENT DURATIONS.
(WEIBULL DISTRIBUTION).

COEFFICIENT/(ASYMPTOTIC STANDARD ERROR)

VARIABLE	FULL SAMPLE (N=519)	BLACK (N=175)	HISPANIC (N=168)	WHITE (N=176)
ALPHA	0.758	0.835	0.743	0.708
	(0.034)	(0.062)	(0.056)	(0.059)
CONST	-1.366	-1.289	-0.108	-2.126
	(0.069)	(1.089)	(1.199)	(1.303)
AGE	-0.107	-0.070	-0.164	-0.059
	(0.040)	(0.073)	(0.066)	(0.077)
AGESQ/100	0.117	0.052	0.201	0.055
	(0.060)	(0.112)	(0.097)	0.114)
EDUCATION	-0.021	-0.061	-0.025	-0.005
	(0.030)	(0.058)	(0.049)	(0.054)
BLACK	0.398			
	(0.141)			
CHICANO	0.289			
	(0.146)			
LOGLIKELIHOOD†	-1684.94	-554.87	-576.18	-551.73
CHI-SQUARED STATISTIC††	32.92	11.75	11.11	4.00
	(0.000004)	(0.0083)	(0.0111)	(0.26)

TABLE 2 - CONTINUED

SPECIFICATION	FULL SAMPLE (N=519)	BLACK (N=175)	HISPANIC (N=168)	WHITE (N=176)
	STATISTIC / (PROB $[\chi^2(k-1) > Z_{(k)})]$			
$Z_{(2)}$	3.62 (0.057)	1.45 (0.228)	2.99 (0.084)	0.444 (0.505)
$Z_{(3)}$	4.49 (0.106)	1.46 (0.478)	3.26 (0.196)	1.257 (0.533)
$Z_{(4)}$	15.28 (0.0016)	5.90 (0.117)	15.05 (0.0018)	1.768 (0.622)

† A chi-squared (likelihood-ratio) test for the hypothesis that the "race-specific" model is better than the "full sample" model takes the value 4.32, which does not fall in the critical region of a $\chi^2(9)$ distribution.

†† Likelihood ratio test for the hypothesis that all the coefficients of the covariates, except that of the constant, are zero. In brackets is the probability that a chi-square variate with appropriate degrees of freedom is greater than the likelihood ratio statistic (degrees of freedom are 5 for full sample and 3 for race specific samples).

TABLE 3

HAZARD FUNCTION ESTIMATES AND TEST STATISTICS.
UNEMPLOYMENT DURATIONS.
(WEIBULL DISTRIBUTION).

COEFFICIENT/(ASYMPTOTIC STANDARD ERROR)

VARIABLE	FULL SAMPLE (N=348)	BLACK (N=114)	HISPANIC (N=121)	WHITE (N=113)
ALPHA	0.938	0.917	1.040	0.918
	(0.044)	(0.076)	(0.088)	(0.073)
CONST	-3.487	-2.932	-5.830	-4.621
	(0.794)	(1.112)	(1.481)	(1.580)
AGE	0.087	0.075	0.128	0.129
	(0.045)	(0.075)	(0.079)	(0.086)
AGESQ/100	-0.145	-0.126	-0.188	-0.210
	(0.069)	(0.114)	(0.119)	(0.130)
EDUCATION	-0.005	-0.094	0.082	0.045
	(0.034)	(0.057)	(0.061)	(0.065)
BLACK	-0.705			
	(0.156)			
CHICANO	-0.376			
	(0.158)			
LOGLIKELIHOOD†	-996.79	-341.57	-327.44	-323.73
CHI-SQUARED STATISTIC††	27.26	2.94	5.86	4.86
	(0.000051)	(0.40)	(0.12)	(0.18)

TABLE 3 - CONTINUED

	SPECIFICATION	TEST		STATISTICS
	FULL SAMPLE (N=348)	BLACK (N=114)	HISPANIC (N=121)	WHITE (N=113)
	STATISTIC / (PROB $[\chi^2(k-1) > Z_{(k)})]$			
$Z_{(2)}$	6.84 (0.0089)	2.21 (0.137)	2.02 (0.155)	3.01 (0.083)
$Z_{(3)}$	6.84 (0.0327)	2.45 (0.294)	2.20 (0.332)	3.12 (0.210)
$Z_{(4)}$	12.36 (0.0062)	3.22 (0.359)	4.08 (0.253)	4.35 (0.226)

† A chi-squared (likelihood-ratio) test for the hypothesis that the "race-specific" model is better than the "full sample" model takes the value 8.10, which does not fall in the critical region of a $\chi^2(9)$ distribution.

†† Likelihood ratio statistic for the hypothesis that all the coefficients of the covariates, except that of the constant, are zero. In brackets is the probability that a chi-square variate with appropriate degrees of freedom is greater than the likelihood ratio statistic (degrees of freedom are 5 for full sample and 3 for race specific samples).

TABLE 4

HAZARD FUNCTION ESTIMATES AND TEST STATISTICS.
UNEMPLOYMENT DURATIONS.
(EXPONENTIAL DISTRIBUTION).

COEFFICIENT/(ASYMPTOTIC STANDARD ERROR)

VARIABLE	FULL SAMPLE (N=348)	BLACK (N=114)	HISPANIC (N=121)	WHITE (N=113)
CONST	-3.673 (0.784)	-3.079 (1.085)	-5.652 (1.429)	-5.087 (1.535)
AGE	0.088 (0.045)	0.070 (0.074)	0.126 (0.079)	0.139 (0.086)
AGESQ/100	-0.147 (0.069)	-0.119 (0.112)	-0.184 (0.118)	-0.227 (0.131)
EDUCATION	-0.004 (0.034)	-0.099 (0.056)	0.080 (0.061)	0.053 (0.065)
BLACK	-0.745 (0.153)			
HISPANIC	-0.382 (0.158)			
LOGLIKELIHOOD†	-997.75	-342.16	-327.55	-324.34
CHI-SQUARED STATISTIC††	30.58 (0.000011)	3.12 (0.37)	5.68 (0.13)	5.84 (0.12)

TABLE 4 - CONTINUED

	SPECIFICATION TEST STATISTICS			
	FULL SAMPLE (N=348)	BLACK (N=114)	HISPANIC (N=121)	WHITE (N=113)
	STATISTIC / (PROB $[\chi^2(k-1) > Z_{(k)})]$			
$Z_{(2)}$	10.17 (0.0014)	3.48 (0.062)	1.40 (0.237)	6.79 (0.0092)
$Z_{(3)}$	11.68 (0.0029)	3.82 (0.148)	2.24 (0.148)	7.24 (0.0268)
$Z_{(4)}$	13.69 (0.0034)	3.97 (0.265)	4.33 (0.228)	7.25 (0.0643)

† A chi-squared (likelihood-ratio) test for the hypothesis that the "race-specific" model is better than the "full sample" model takes the value 7.4, which does not fall in the critical region of a $\chi^2(9)$ distribution.

†† Likelihood ratio statistic for the hypothesis that all the coefficients of the covariates, except that of the constant, are zero. In brackets is the probability that a chi-square variate with appropriate degrees of freedom is greater than the likelihood ratio statistic (degrees of freedom are 5 for full sample and 3 for race specific samples).

acceptable for the Black and Hispanic sample. Further, a comparison of the diagnostics in Table 3 and Table 4 shows that the Weibull distribution which is more "flexible" than the Exponential provides a better fit for the data, in all the three race categories.

We now show that a visual check of our distributional assumptions using generalized residual plots, reinforces the conclusions obtained by using the diagnostic statistics. The definition of generalized residuals used here is that given by Cox and Snell (1968). Consider the general model

$$Y_i = g_i(X_i, \beta, \epsilon_i) \qquad i = 1, 2, \ldots, n \qquad \text{—— (62)}$$

where $Y = (Y_i)$ is a vector random variable, X_i an observed characteristic vector for the i^{th} observation, β a vector of unknown parameters and $\epsilon = (\epsilon_i)$ a vector of independent and identically distributed unobserved random variables. Let $\hat{\beta}$ be the Maximum Likelihood estimate of β. If the equation

$$Y_i = g_i\,(X_i, \hat{\beta}, \hat{\epsilon}_i) \qquad \text{—— (63)}$$

has a unique solution for $\hat{\epsilon}_i$, namely

$$\hat{\epsilon}_i = h_i\,(Y_i, X_i, \hat{\beta}) \qquad \text{—— (64)}$$

then $\hat{\epsilon}_i$ is called the generalized residual corresponding to Y_i and model specified in (62). The generalized residuals from a fitted model essentially mean a set of transformations of the observations that have a know simple distribution (at least in large samples) when the assumed model is correct, but a different distribution when it is not.

Consider the hazard function $\pi(t,X,\beta)$ where X is a vector of co-variates and β a vector of unknown parameters, of an arbitrary distribution. Define the integrated hazard as

$$\epsilon(t,X) = \int_0^t \pi(s,X,\beta)\ ds \qquad (65)$$

It is easy to see that $\epsilon(t,X)$ has a unit Exponential distribution since

$$e^{-\epsilon(t,X)} = \text{Prob}\ [\tilde{t} > t] = \text{Prob}\ [\epsilon(\tilde{t},X) > \epsilon(t,x)] \qquad (66)$$

Hence, minus the logarithm of the survivor function should plot on a 45° line through the origin. In the case of the Weibull distribution under proportional hazard specification, the generalized residuals are defined by

$$\hat{\epsilon}(t,X) = \int_0^t \hat{\alpha}\ s^{\hat{\alpha}-1}\ \exp\ (X\beta)\ ds = t^{\hat{\alpha}}\ \exp\ (X\beta) \qquad (67)$$

In general we have both complete and censored durations. Let $\tilde{t}$ and $\tilde{t}^c$ be independent random variables where $\tilde{t}$ is the duration of an event and $\tilde{t}^c$ a censoring time. The observed duration is

$$\tilde{t}^0 = \text{Min}\ [\tilde{t},\ \tilde{t}^c] \qquad (68)$$

With censoring we can only calculate $\epsilon(\tilde{t}^0, X)$ which does not have a unit Exponential distribution. In the case of right censored observations we can use the Kaplan-Meier procedure[14] to estimate the survivor function of $\epsilon(\tilde{t},X)$ from $\epsilon(\tilde{t}^0,X)$ data and this is distributed unit exponential when the model is right. Consider the ordered sequence of $\epsilon(t_i^0, X_i)$, $i = 1,2,\ldots,N$. For each $\epsilon(t_i^0, X_i)$ corresponding to an uncensored observation we estimate the hazard function at the value of $\epsilon(t_i^0, X_i)$ as the ratio of the number of residuals with value equal to the particular $\epsilon(t_i^0, X_i)$ and the number of residuals greater than or equal to it. Let this ratio for the j^{th} ordered uncensored residual be m_j. Then, the Kaplan-Meier estimate of the survivor

---

[14]see Kaplan and Meier (1958), Cox and Oakes (1984) and Chesher and Lancaster (1985).

function of the residual at value a is

$$\prod_{j=1}^{a} (1 - m_j) \qquad (69)$$

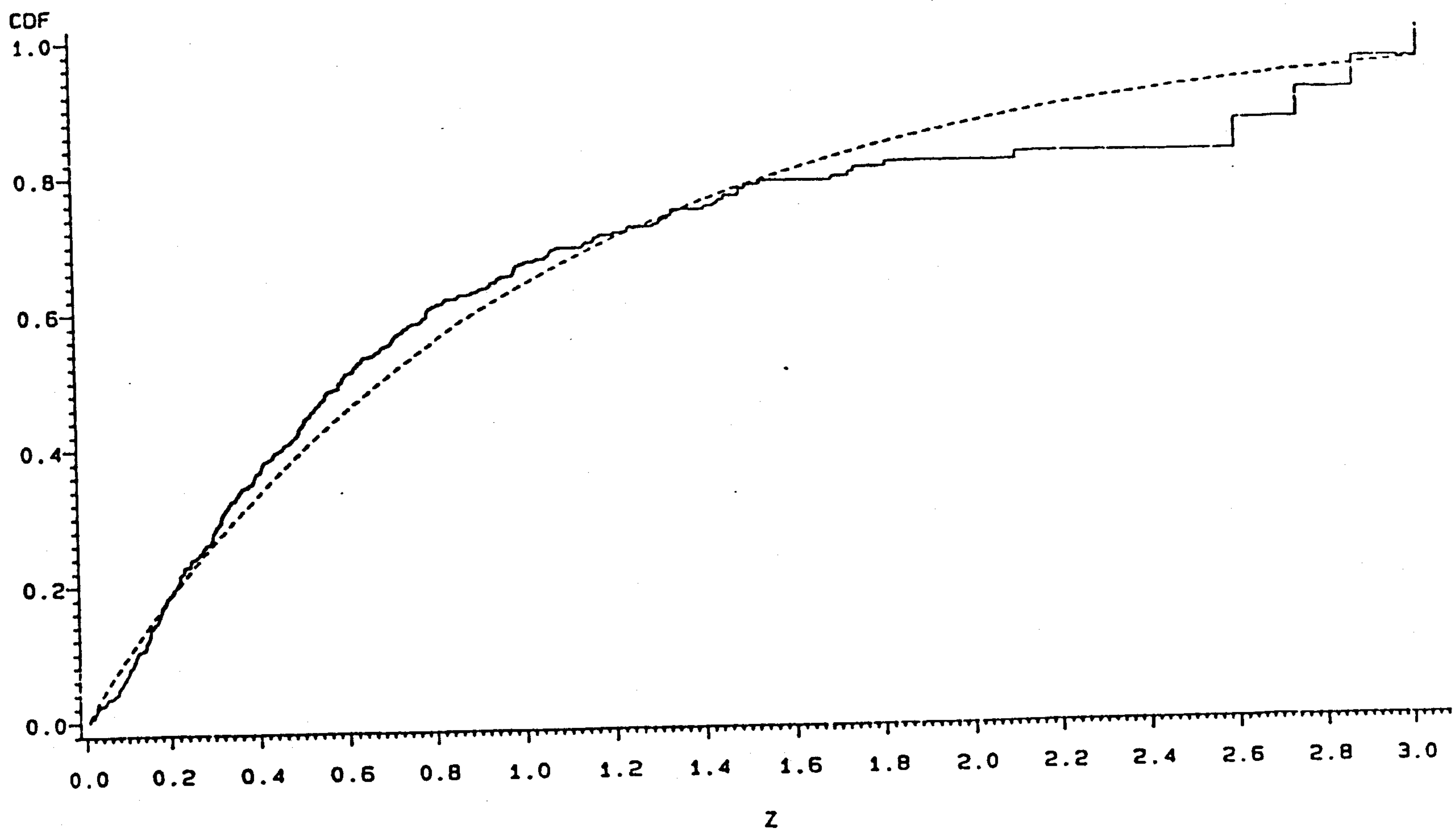
and minus the logarithm of the survivor function is given by

$$-\sum_{j=1}^{a} \ln (1 - m_j) \approx \sum_{j=1}^{a} m_j \qquad (70)$$

The plot of minus the logarithm of the residual survivor function should give a 45° line through the origin in large samples, when the model is right.

Figures 1 to 8 depict the plots for the employment durations when the distribution is assumed Weibull. As the diagnostics brought out, the plots show that the Weibull does a reasonable job for the White and Black sample but not very well for the Hispanic sample and the full sample.

GENERALIZED RESIDUAL PLOTS.
FULL SAMPLE. WEIBULL DISTRIBUTION.



--------- CDF OF GENERALIZED RESIDUALS
- - - - - CDF OF UNIT EXPONENTIAL DISTRIBUTION

FIGURE 1

GENERALIZED RESIDUAL PLOTS.
FULL SAMPLE. WEIBULL DISTRIBUTION.

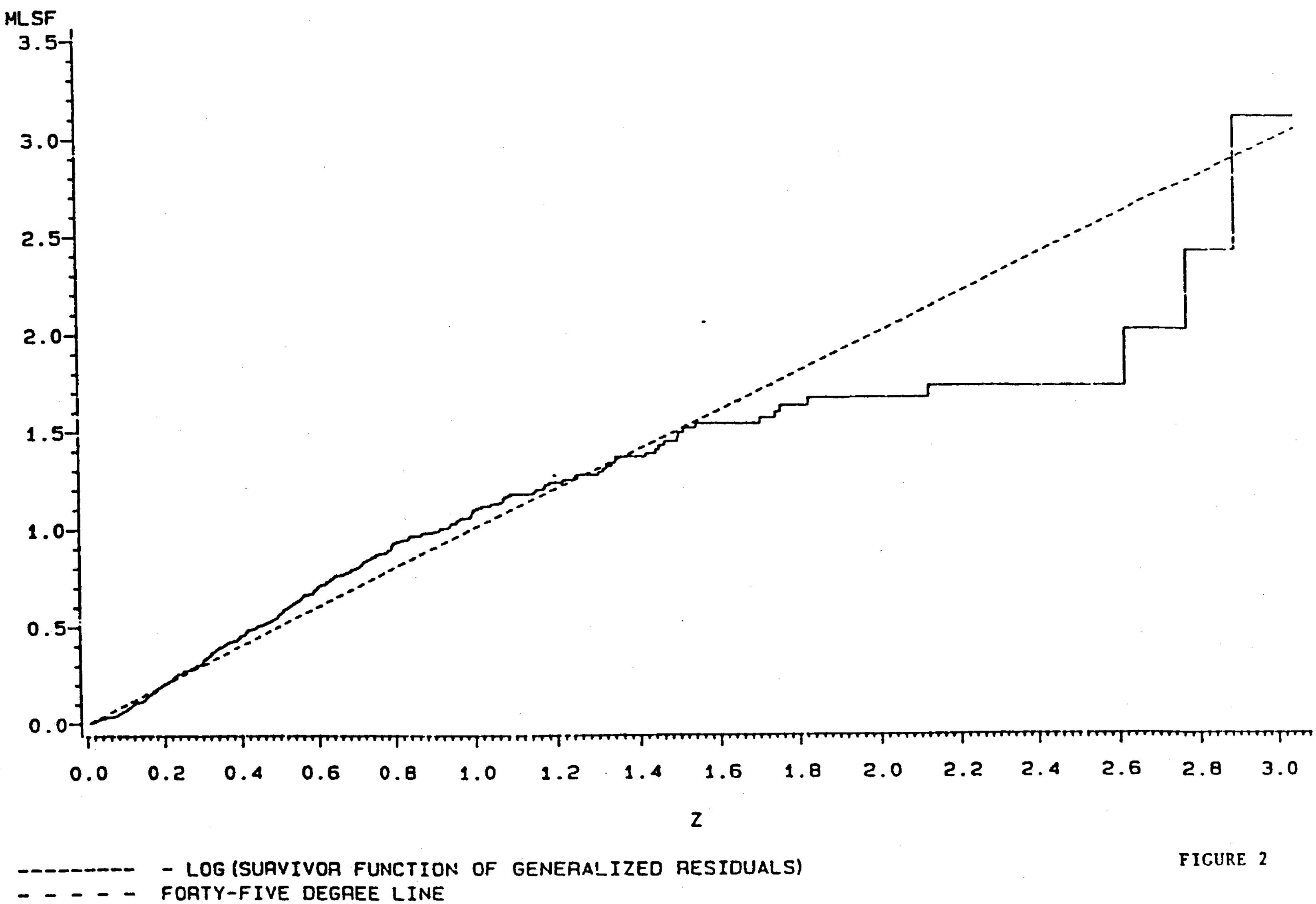


FIGURE 2

GENERALIZED RESIDUAL PLOTS.
RACE=BLACK. WEIBULL DISTRIBUTION.

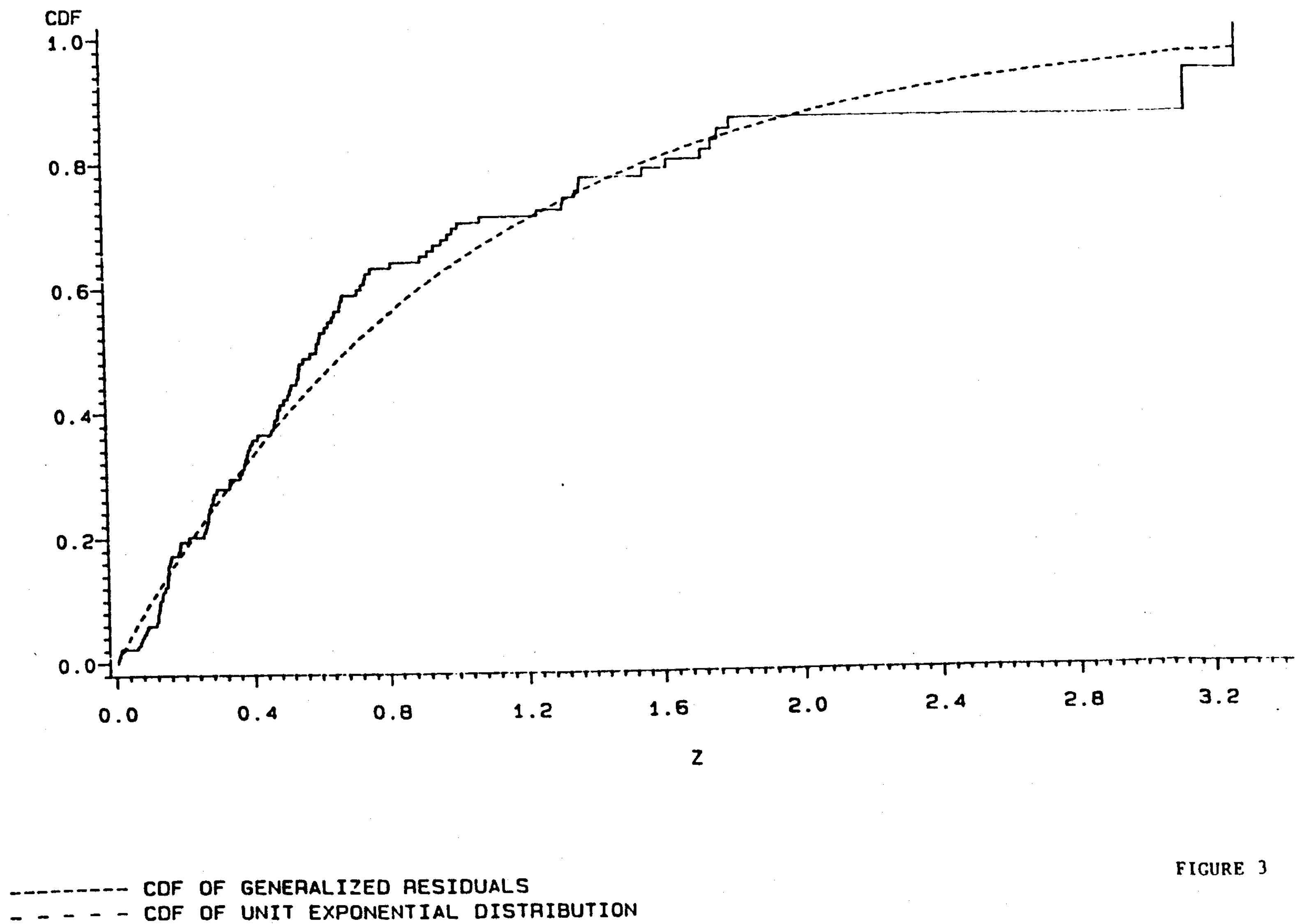


FIGURE 3

GENERALIZED RESIDUAL PLOTS.
RACE=BLACK. WEIBULL DISTRIBUTION.

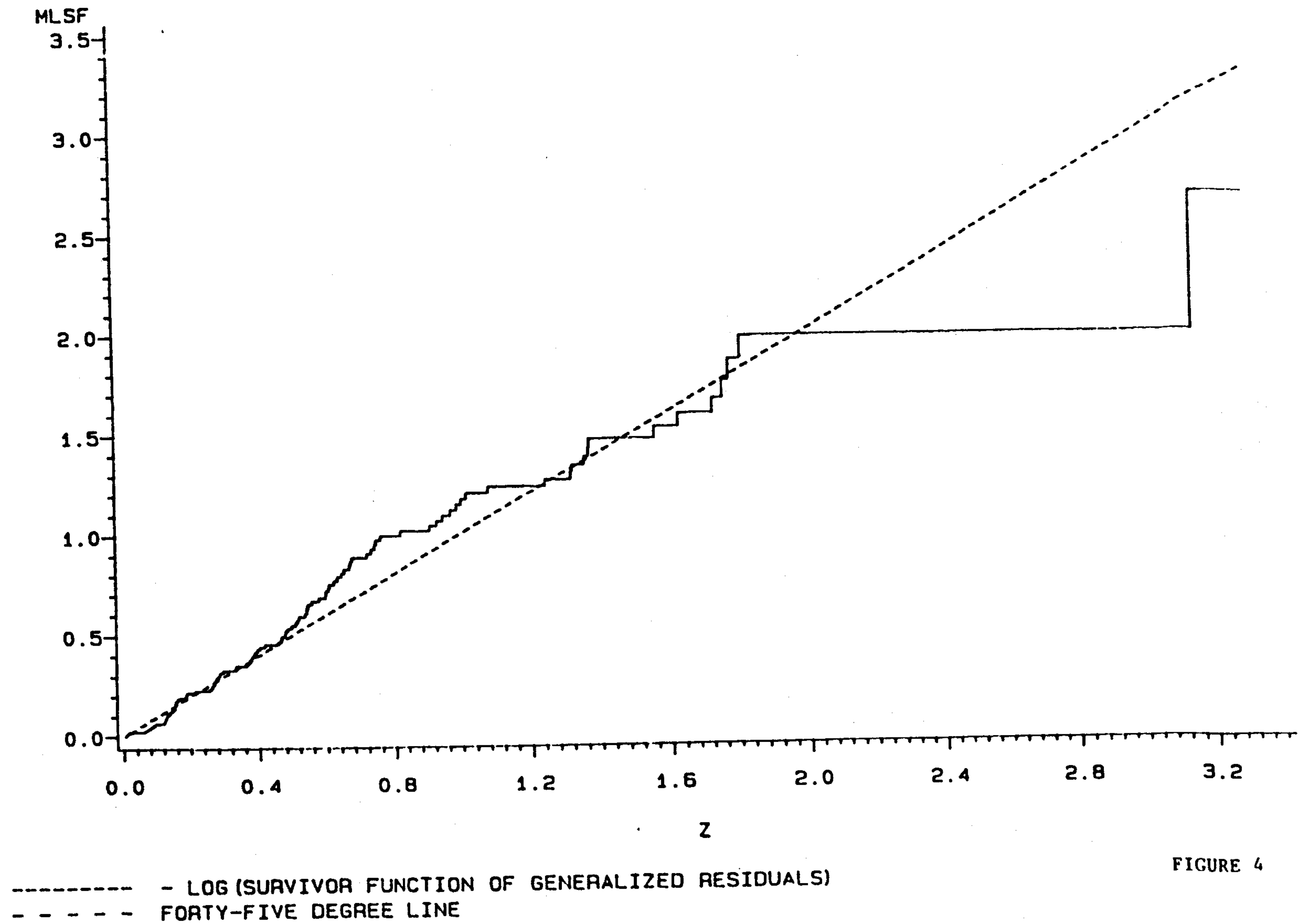


FIGURE 4

GENERALIZED RESIDUAL PLOTS.
RACE=HISPANIC. WEIBULL DISTRIBUTION.

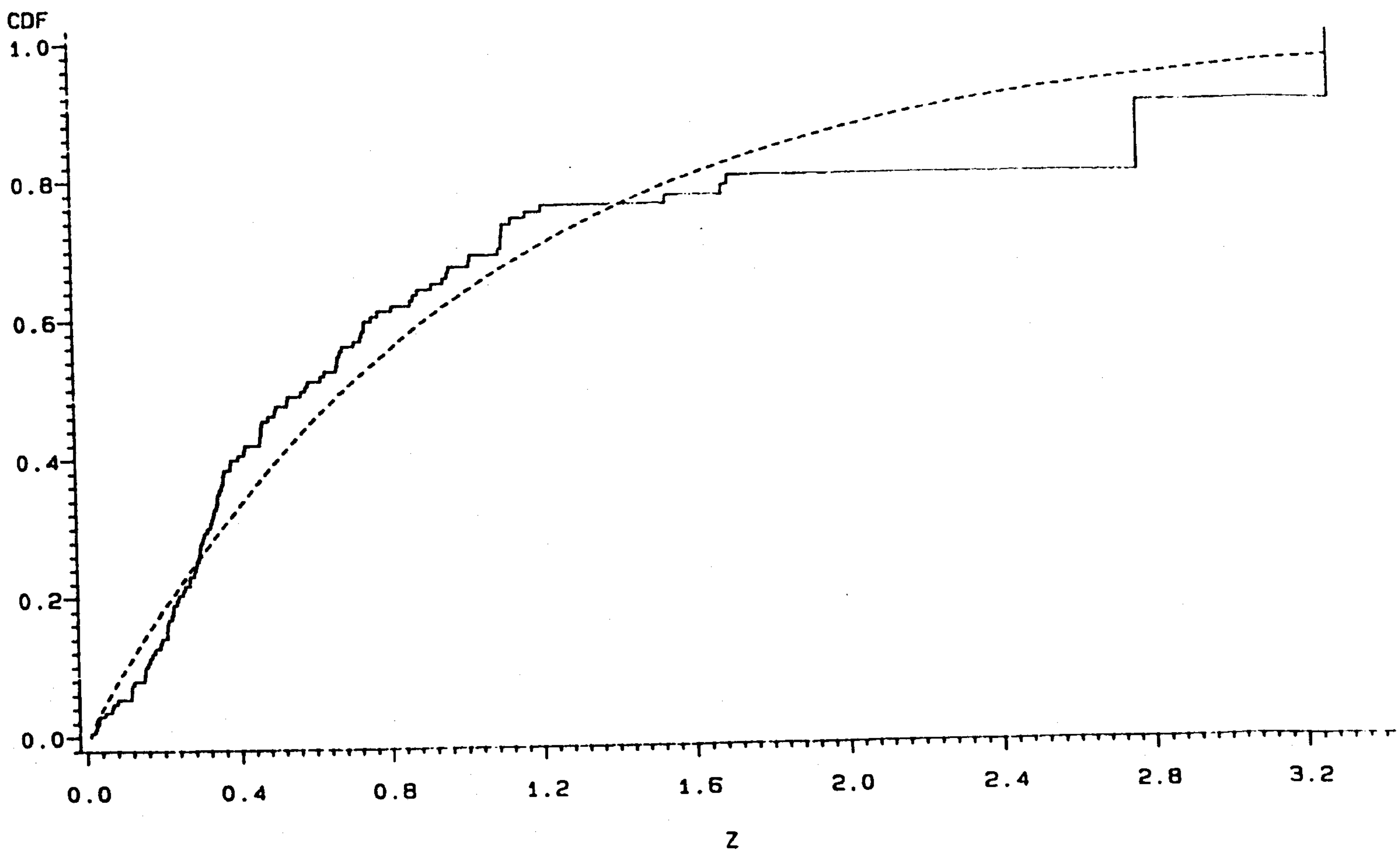


FIGURE 5

--------- CDF OF GENERALIZED RESIDUALS
- - - - - CDF OF UNIT EXPONENTIAL DISTRIBUTION

GENERALIZED RESIDUAL PLOTS.
RACE=HISPANIC. WEIBULL DISTRIBUTION.

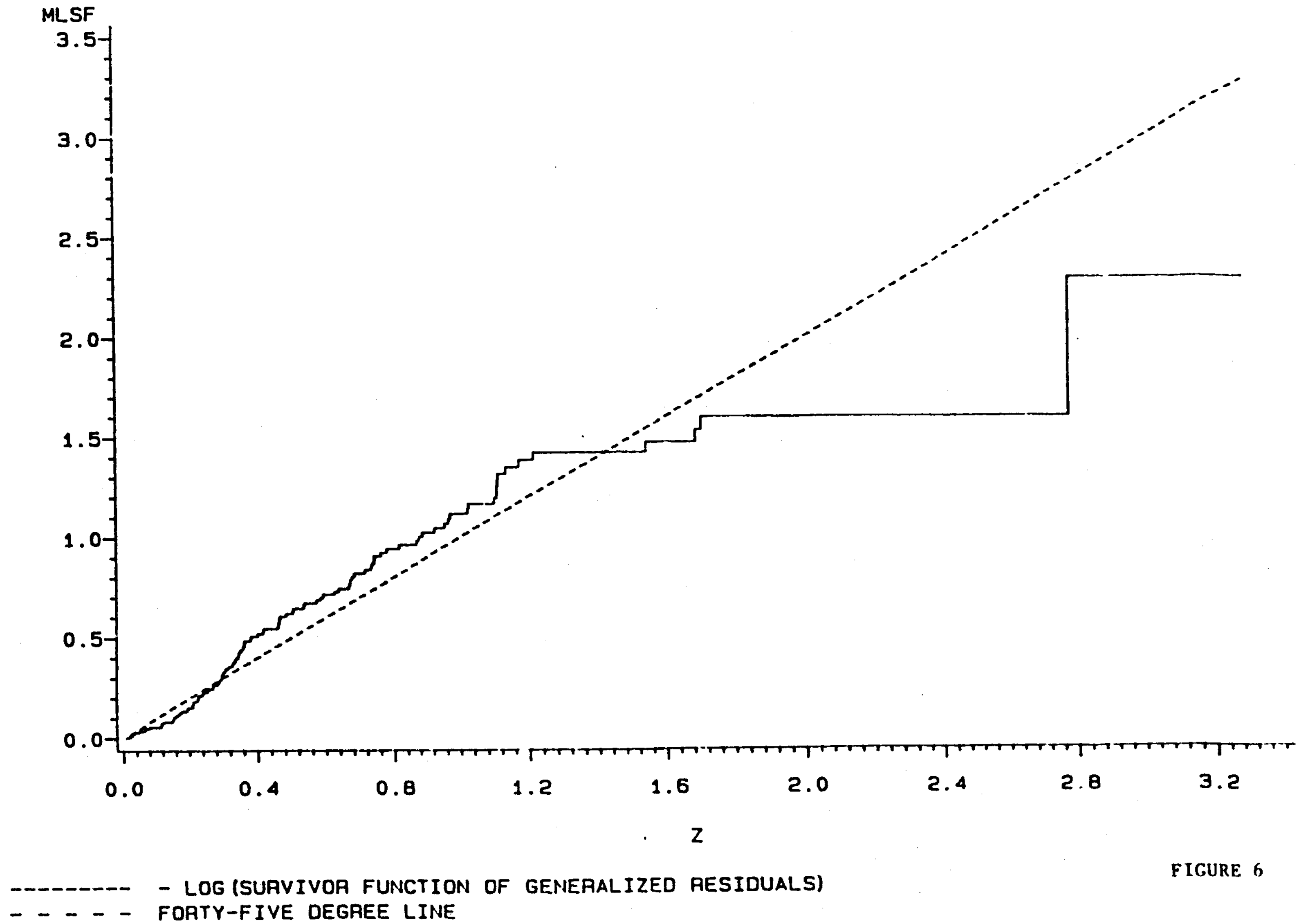


FIGURE 6

GENERALIZED RESIDUAL PLOTS.
RACE=WHITE. WEIBULL DISTRIBUTION.

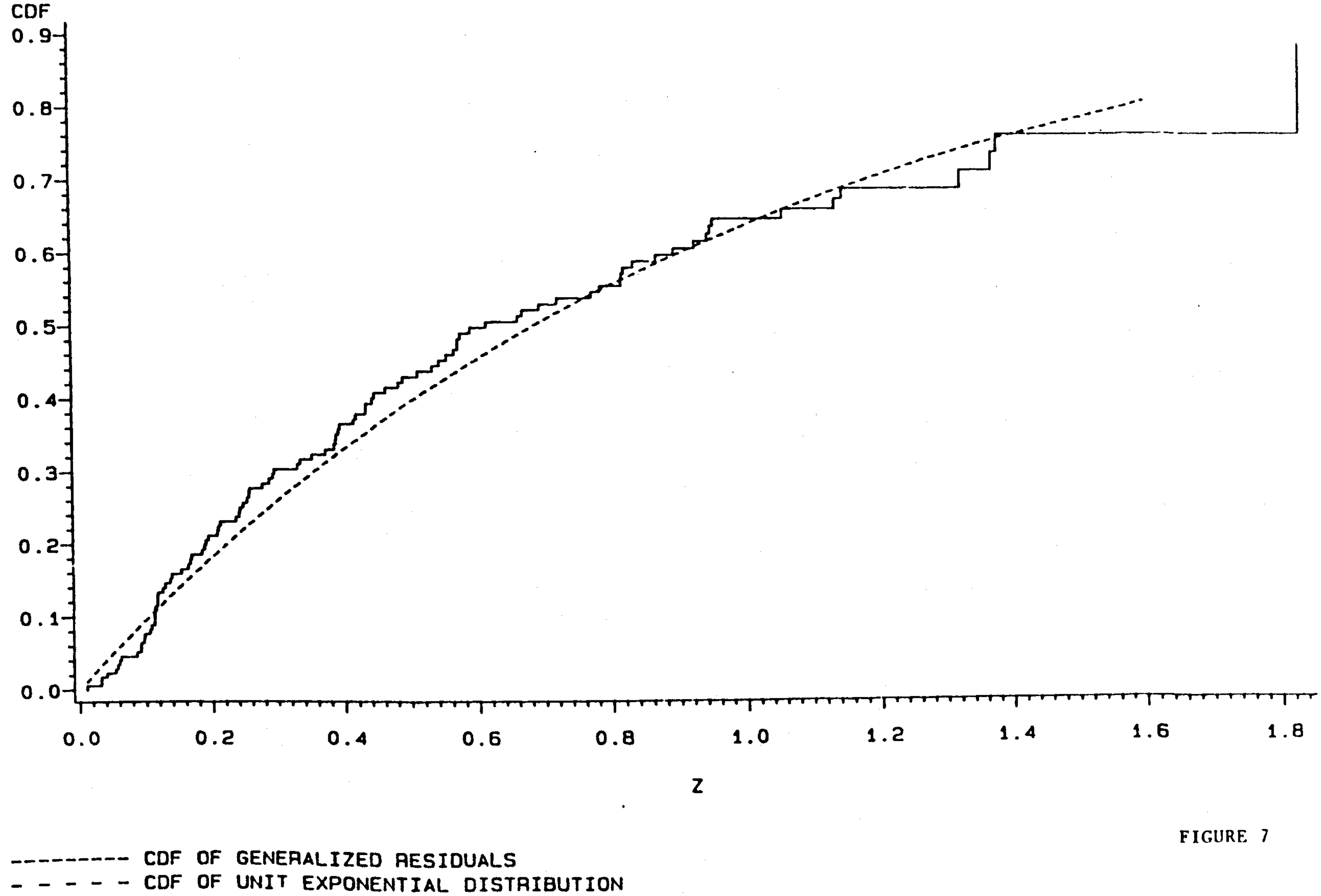


FIGURE 7

GENERALIZED RESIDUAL PLOTS.
RACE=WHITE. WEIBULL DISTRIBUTION.

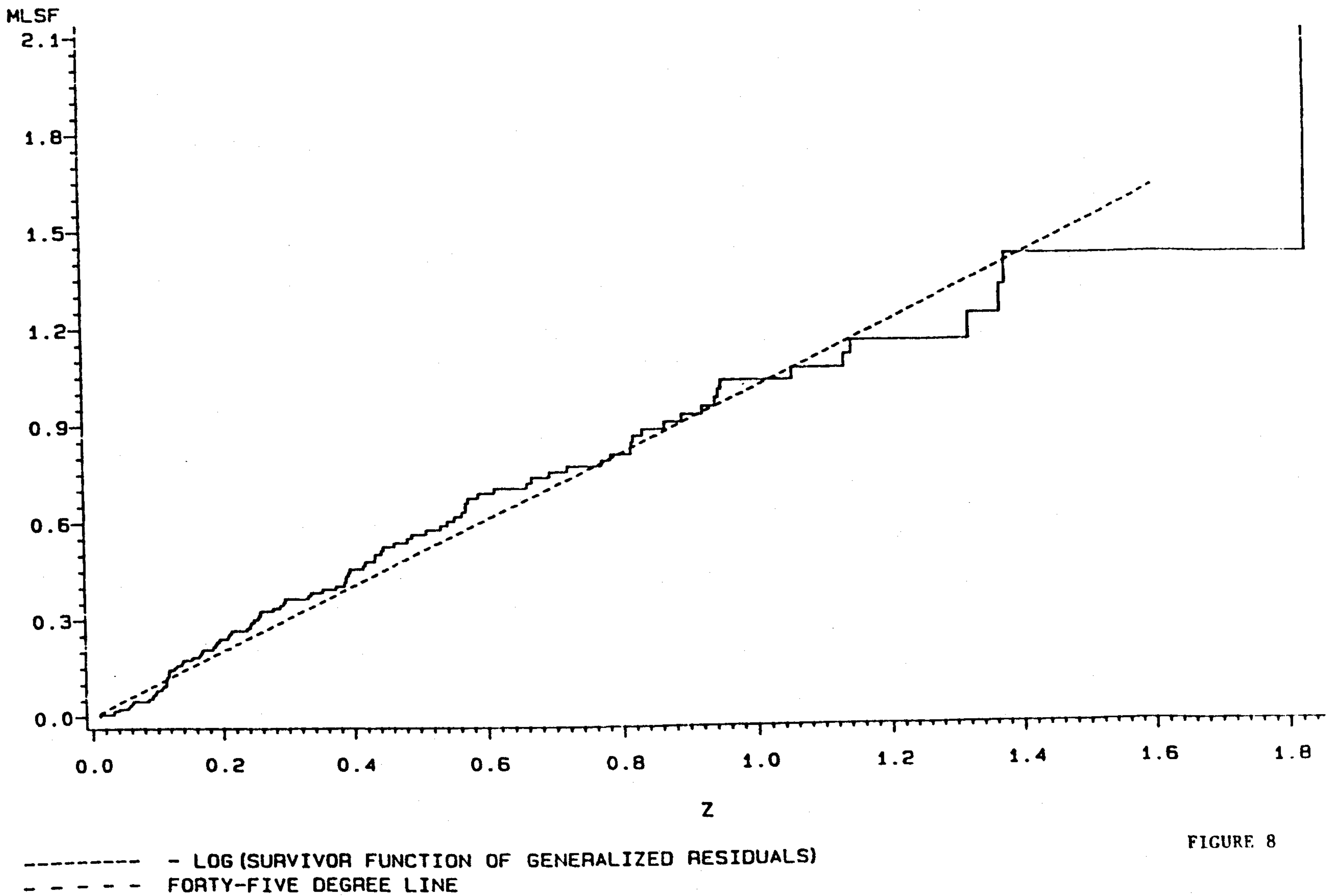


FIGURE 8

CONCLUSION

The diagnostics for assessing the appropriateness of the distributional specification were based on the idea that if the empirical distribution function (conditional on covariates) is "close" to that of the Weibull, then it should satisfy the restrictions on moments implied by the Weibull distribution. The diagnostics are easy to compute and provide information on goodness of fit in addition to that provided by likelihood ratio tests. In the case of a bad fit the analyst's strategy should be to include other covariates which he/she thinks may influence the durations being examined. Alternatively, estimation under other distributional assumptions maybe done. It should be mentioned that diagnostics of this nature do not provide evidence of any particular kind of misspecification. Rather they indicate whether the estimated model is appropriate or not.

It was shown that two ways of developing diagnostics for uncorrected heterogeneity based on approximations to the distribution of the heterogeneity component lead to essentially the same statistics. Further, these diagnostics for heterogeneity were equivalent to testing whether a particular moment restriction implied by the null was satisfied. Unlike the linear regression model where omitted regressors create a problem only if the left out variables are correlated with those included, in the duration models we are considering, omitting regressors, even if they are uncorrelated with those included, can lead to biased estimates of the coefficients of the (included) covariates. Consequently, it maybe important to look at (heterogeneity) diagnostics even in the absence of tightly specified alternatives.

APPENDIX A

PROOF OF THEOREM:

Consider for $|a| < 1$,

$$(1 - a)^{-1} \exp\left\{ - \frac{a\, t^{\alpha}}{(1 - a)} \right\}$$

$$= (1 - a)^{-1} \sum_{k=0}^{\infty} \frac{(-t^{\alpha})^k}{k!} \frac{a^k}{(1 - a)^k}$$

$$= \sum_{k=0}^{\infty} \frac{(-t^{\alpha})^k}{k!} \left[a^k \sum_{m=0}^{\infty} \binom{k + m}{k} a^m \right]$$

$$= \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} \left[\binom{k + m}{k} \frac{(-t^{\alpha})^k}{k!} a^{m+k} \right]$$

$$= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \binom{n}{k} \frac{(-t^{\alpha})^k}{k!} a^n$$

$$= \sum_{n=0}^{\infty} L_n(t)\, a^n \qquad \text{—— (1)}$$

Let $G(t,a) = (1 - a)^{-1} \exp\left\{ \frac{-a\, t^{\alpha}}{(1 - a)} \right\}$

$$= \sum_{n=0}^{\infty} L_n(t)\, a^n \qquad \text{—— (2)}$$

Then, for $|a| < 1$ and $|b| < 1$, we have

$$G(t,a)\; G(t,b) = [(1-a)(1-b)]^{-1} \exp\left\{ -t^{\alpha} \left[\frac{a}{(1-a)} + \frac{b}{(1-b)} \right] \right\} \qquad \text{—— (3)}$$

Therefore,

$$\int_0^\infty G(t,a) \quad G(t,b) \quad \alpha \; t^{\alpha-1} \quad \exp \quad (-t^\alpha) \quad dt$$

$$= \; [(1-a)(1-b)]^{-1} \quad \int_0^\infty \alpha \; t^{\alpha-1} \quad \exp \; \left\{ -t^\alpha \; \left[\frac{1-ab}{(1-a)(1-b)} \right] \right\} dt \qquad \text{—— (4)}$$

$$= \; \frac{1}{(1 \; - \; ab)}$$

Now the product in (3) is also equal to

$$G(t,a) \quad G(t,b) \; = \; \left(\sum_{n=0}^{\infty} \; L_n \; (t) \quad a^n \right) \; \left(\sum_{m=0}^{\infty} \; L_m \quad (t) \quad b^m \right)$$

$$= \; \sum_{n=0}^{\infty} \quad \sum_{m=0}^{\infty} \; L_n \; (t) \quad L_m \; (t) \quad a^n \, b^m \qquad \text{—— (5)}$$

Integrating (5) we get

$$\int_0^\infty \; G(t,a) \quad G(t,b) \quad \alpha t^{\alpha-1} \quad \exp \quad (-t^\alpha) \quad dt$$

$$= \; \sum_{n=0}^{\infty} \quad \sum_{m=0}^{\infty} \; \left[\; \int_0^\infty \; L_n(t) \quad L_m(t) \quad \alpha t^{\alpha-1} \quad \exp \quad (-t^\alpha) \; dt \right]. \quad a^n \, b^m$$

$$\text{—— (6)}$$

(Fubini's theorem justifies interchange of the integral and summations).

Also, from (4) we have

$$\int_0^\infty G(t,a) \quad G(t,b) \quad \alpha t^{\alpha-1} \quad \exp \; (-t^a) \quad dt \; = \; \sum_{n=0}^{\infty} \; a^n \, b^n \qquad \text{—— (7)}$$

Comparing coefficients of $a^n \, b^m$ in (6) and (7) we get

$$\int_0^\infty \; L_n \; (t) \quad L_m \; (t) \quad \alpha t^{\alpha-1} \quad \exp \; (-t^\alpha) \quad dt \; = \; \delta_{mn} \qquad m,n \; = \; 0,1,2,\ldots$$

the desired result.

APPENDIX B

In this appendix we consider the case when some of the duration spells are censored. Suppose the data $\{t_i, X_i, \delta_i\}$ $i = 1,2,\ldots,N$ comprises of complete and right censored spells. Then, N^{-1} times the loglikelihood function can be written as

$$l = N^{-1} \sum_{i=1}^{N} [\ \delta_i \ln p_i^*(t_i) + (1 - \delta_i) \ln \bar{P}_i^* (t_i)\ ] \qquad (1)$$

In the above equation, $\bar{P}*(t)$ is the survivor function corresponding to the approximate density $p*(t)$, and is given by

$$\bar{P}* (t) = \int_t^\infty p*(u)\ du$$

$$\int_t^\infty \alpha u^{\alpha-1} \eta \exp(-\eta u^\alpha)\ du$$

$$+ \theta_2 \int_t^\infty L_2 (\eta^{\frac{1}{\alpha}} u)\ \alpha u^{\alpha-1} \eta \exp(-\eta u^\alpha)\ du$$

$$+\ldots+ \theta_n \int_t^\infty L_n (\eta^{\frac{1}{\alpha}} u)\ \alpha u^{\alpha-1} \eta \exp(-\eta u^\alpha)\ du$$

$$= \exp(-\eta t^\alpha) + \sum_{j=2}^{N} \theta_j \mathfrak{L}_j (\eta,t) \qquad (2)$$

where

$$\mathfrak{L}_j (\eta,t) = \int_t^\infty L_j (\eta^{\frac{1}{\alpha}} u)\ \alpha u^{\alpha-1} \exp(-\eta u^\alpha)\ du \qquad (3)$$

For example,

$$\mathcal{L}_2 \quad (\eta, t) \quad = \quad \frac{1}{2} \int_t^\infty [\eta^2 u^{2\alpha} - 4\eta u^\alpha + 2] \quad \alpha u^{\alpha-1} \eta \quad \exp(-\eta u^\alpha) \quad du$$

$$= \frac{1}{2} \Gamma(3, \eta t^\alpha) - 2 \quad \Gamma(2, \eta t^\alpha) + \Gamma(1, \eta t^\alpha) \qquad \text{—— (4)}$$

where

$$\Gamma(a,b) \quad = \quad \int_b^\infty u^{a-1} e^{-u} \quad du$$ is the incomplete gamma function.[1] —— (5)

Similarly, we get

$$\mathcal{L}_3 \; (\eta, t) \quad = \quad \frac{1}{6} \; \Gamma(4, \eta t^\alpha) + \frac{3}{2} \Gamma(3, \eta t^\alpha) - 3 \; \Gamma(2, \eta t^\alpha) + \Gamma(1, \eta t^\alpha)$$

$$\mathcal{L}_4 \; (\eta, t) \quad = \quad \frac{1}{24} \Gamma(5, \eta t^\alpha) - \frac{2}{3} \Gamma(4, \eta t^\alpha) + 3 \; \Gamma(3, \eta t^\alpha)$$

$$- 4 \; \Gamma(2, \eta t^\alpha) + \Gamma(1, \eta t^\alpha) \qquad \text{—— (6)}$$

Consider testing $\theta_2 = 0$. Differentiating (1) and evaluating under the null, gives

$$\frac{\delta 1}{\delta \theta_2} \quad = \quad N^{-1} \sum_{i=1}^{N} \left[\delta_i \; L_2 \; (\eta_i^{\frac{1}{\alpha}} \; t_i) \right.$$

$$\left. + (1 - \delta_i) \quad \exp(\eta_i t_i^\alpha) \quad \mathcal{L}_2 \; (\eta_i, t_i) \right]$$

—— (7)

The Lagrange multiplier statistic is

$$z_{(2)} \quad = \quad \frac{\left\{ \frac{\delta 1}{\delta \theta_2} \right\}^2}{V_{(2)}} \qquad \text{—— (8)}$$

---

[1]see Abramovitz and Stegun (1972), pg. 260.

where $V_{(2)}$ is an estimate of the variance of $\frac{\delta l}{\delta \theta_2}$, calculated in a manner analogous to that in equation (27).

For higher order tests, a statistic similar to $Z_{(n)}$ in (29) is calculated. Note that the calculation of $\hat{L}_j(\eta, t)$ is quite straight forward. The incomplete gamma function can be expressed as

$$\Gamma(a,b) = \Gamma(a) \quad \text{Prob}\,\{\chi^2(2a) > 2b\} \qquad (9)$$

and since most computer packages (e.g., SAS) calculate gamma functions and chi-square integrals the required calculations for above derived statistics are easy.

APPENDIX C

The loglikelihood function based on complete and right censored durations can be written as

$$l^* = \sum_{i=1}^{N} [\delta_i \ln f_i^* (t_i, X_i) + (1 - \delta_i) \ln \bar{F}_i^* (t_i, X_i)] \quad \text{——} (1)$$

where $\bar{F}^*(t,X)$ is the survivor function corresponding to the density $f^*(t,X)$. We know the survivor function corresponding to the density in (52) is

$$\bar{F}^* (t,X) = \bar{F} (t,X,u=1) [1 + \frac{\sigma^2}{2} \epsilon^2 (t,X)] \quad \text{——} (2)$$

and it is easy to show that the survivor function corresponding to the density in (45) is

$$\bar{F}^* (t,X) = \bar{F} (t,X,\nu=0) \left[1 + \frac{\sigma^2}{2} \left\{ \epsilon^2 (t,X) - \epsilon (t,X) \right\} \right] \quad \text{——} (3)$$

Using (2), the score for σ^2 under the null is

$$S = \frac{1}{2N} \sum_{i=1}^{N} [\delta_i (\epsilon_i^2 - 2 \epsilon_i) + (1 - \delta_i) \epsilon_i^2] \quad \text{——} (4)$$

and using (3) the score for σ^2 under the null is

$$S = \frac{1}{2N} \sum_{i=1}^{N} [\delta_i (1 - 3 \epsilon_i + \epsilon_i^2) + (1 - \delta_i) (\epsilon_i^2 - \epsilon_i)] \quad \text{——} (5)$$

The diagnostic statistics when censored observations are present are based on the scores defined in (4) and (5).

REFERENCES

1. Abramowitz, M. and I.A. Stegun, eds., [1972]: Handbook of Mathematical Functions, Dover, New York.

2. Breslow, D., [1975]: "Analysis of Survival Data under the Proportional Hazard Model," International Statistical Review, 43, 45-58.

3. Breusch, T.S. and A.R. Pagan, [1980]: "The Lagrange Multiplier Test and its Applications to Model Specification in Econometrics," Review of Economic Studies, XLVII, 239-253.

4. Burdett, K., N.M. Kiefer, D.T. Mortensen and G.R. Neumann, [1984]: "Earnings, Unemployment, and the Allocation of Time over Time," Review of Economic Studies, 559-578.

5. Burdett, K., N.M. Kiefer, and S. Sharma, [1985]: "Layoffs and Duration Dependence in a Model of Turnover," Journal of Econometrics, Annals, 28, 51-70.

6. Chesher, A., [1984]: "Testing for Neglected Heterogeneity," Econometrica, 52, 865-872.

7. Chesher, A., and T. Lancaster, [1985]: "Residuals, Tests and Plots with a Job Matching Illustration," mimeo.

8. Chihara, T.S., [1978]: An Introduction to Orthogonal Polynomials, Gordon and Breach, Science Publishers.

9. Cox, D.R., [1962]: Renewal Theory, Methuen, London.

10. Cox, D.R., [1972]: "Regression Models and Life Tables," Journal of the Royal Statistical Society, B, 34, 187-220.

11. Cox, D.R., [1983]: "Some Remarks on Overdispersion," Biometrica, 70 269-274.

12. Cox, D.R., and D. Oakes, [1984]: Analysis of Survival Data, Chapman and Hall, London.

13. Cox, D.R., and E.J. Snell, [1968]: "A General Definition of Residuals," Journal of the Royal Statistical Society, B, 2, 248-275.

14. Engle, R.F., [1982]: "A General Approach to Lagrange Multiplier Model Diagnostics," Journal of Econometrics, 20, 83-104.

15. Feigl, P. and M. Zelen, [1965]: "Estimation of Exponential Survival Probabilities with Concomitant Information," Biometrics, 21, 826-838.

16. Flinn, C. and J.J. Heckman, [1982]: "Models for the Analysis of Labor Force Dynamics," in Advance in Econometrics, vol. I, JAI Press.

17. Greenberg, R.A., S. Bayard, and D. Byar, [1974]: "Selecting Concomitant Variables Using a Likelihood Ratio Step-Down Procedure and a Method for Testing Goodness of Fit in an Exponential Survival Model," Biometrics, 30, 601-608.

18. Heckman, J. and B. Singer, [1984]: "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," Econometrica, 52, 271-320.

19. Holford, T.R., [1976]: "Life Tables with Concomitant Information," Biometrics, 32, 587-598.

20. Horvath, W.J., [1968]: "A Statistical Model for the Duration of Wars and Strikes," Behavioral Science, 13, 18-28.

21. Kaplan, E.L. and P. Meier, [1985]: "Nonparametric Estimation from Incomplete Observations," Journal of the American Statistical Association, 53, 457-481.

22. Katz, L., [1985]: "Layoffs, Recall and the Duration of Unemployment," mimeo, University of California, Berkeley.

23. Kay, R., [1977]: "Proportional Hazard Regression Models and the Analysis of Censored Survival Data," Journal of the Royal Statistical Society, C, 26, 227-237.

24. Kennan, J., [1985a]: "The Duration of Strikes in U.S. Manufacturing," Journal of Econometrics, 5-28.

25. Kennan, J., [1985b]: "The Economics of Strikes," in O. Ashenfelter and R. Layard, eds, Handbook of Labor Economics, North-Holland.

26. Kiefer, N.M., [1984]: "A Simple Test for Heterogeneity in Exponential Models of Duration," Journal of Labor Economics, vol. 2, no. 4, 539-549.

27. Kiefer, N.M., [1985]: "Specification Diagnostics based on Laguerre Alternatives for Econometric models of Duration," Journal of Econometrics, Annals, 28, 135-154.

28. Lancaster, T., [1979]: "Econometric Methods for the Duration of Unemployment," Econometrica, 47, 939-956.

29. Lancaster, T., [1985]: "Generalized Residuals and Heterogeneous Duration Models: With Applications to the Weibull Model," Journal of Econometrics, Annals, 28, 155-169.

30. Lundberg, S.J., [1981]: "Unemployment and Household Labor Supply," Ph.D Dissertation, Northwestern University, Evanston, Illinois.

31. Newby, M. and J. Winterton, [1983]: "The Duration of Industrial Stoppages," Journal of the Royal Statistical Society, A, 146, Part I, 62-70.

32. Newey, W.K., [1985]: "Maximum Likelihood Specification Testing and Conditional Moment Tests," Econometrica, 53, 1047-1070.

33. Sharma, S. [1986]: "Labor Market Histories," unpublished Ph.D. dissertation, Cornell University.

34. Szego, G., [1975]: Orthogonal Polynominals, 4th edition, American Mathematical Society Colloquium Publications, Vol. XXIII.

35. Tauchen, G.E., [1984]: "Diagnostic Checking and Evaluation of Econometric Models," Working Papers in Economics, 84-01, Duke University.

36. Zippan, C., and P. Armitage, [1966]: "Use of Concomitant Variables and Incomplete Survival Information in the Estimation of an Exponential Survival Parameter," Biometrics, 22, 665-672.